2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Study on View-independent Hand Posture Recognition

Hyoyoung Jang and Zeungnam Bien
Div. of EE, Dept. of EECS, KAIST
373-1 Guseong-dong, Yuseong-gu, Daejeon 305-701, Korea
email: zbien@ee.kaist.ac.kr

***Abstract* – We describe a method for estimating new hand views from a single 2D hand image using decomposed approach with subgroup-based scheme. With this method, we can get the simplicity in the sense of computation by comparing the image with models in the promising subgroup instead of comparing with all models. It shows more effectiveness in recognition by process depend on each subgroup and easy of extension.**

I. INSTRUCTION

With the rapid growth of computer technology, studies on new human-computer interface (HCI) are active.

Although there are many attempts, vision-based HCI methods by gesture command are emerging as one of the alternatives. Using contact-free input devices, i.e., cameras, we can get more comfortable operating environment with relatively low price. Moreover, it is familiar thing in daily life to use gesture as a way of command. As gesture is informative enough so that it can substitute other communication methods (even more effective in some cases where other ways of command – voice, keyboard input, etc. are unavailable). Nowadays hand gesture recognition is highlighted in relation to the recent results of virtual reality (VR). Comparing with other means of command, hand gesture is the most suitable for representation of 3-dimensional (3D) information.

As a previous research, we have presented a hand-signal recognition method in 3D space [1]. In the method, we get approximate hand model parameter estimates through comparison with models in the predefined hand-signal database. The comparison is performed by template matching. This approach shows structural simplicity and easy of computation.

During the process, not using specific marker devised to classify each fingers, silhouette is the most important information. However hand is highly articulated object and weak for occlusion.

In this paper, we propose hand posture recognition by subgroup based classification as an alternative.

In chapter 2, I will briefly introduce the method proposed in [1] and in chapter 3 suggest subgroup based classification method. Thereafter I will show some experimental results (chap. 4) and bring to a conclusion (chap. 5).

II. PREVIOUS SYSTEM

To begin with, I would like to show the previous system dealt in [1] for clear understanding. The objective of the system is developing hand-signal recognition system assuming car environment. System can be roughly divided into 3 parts – (1) hand region segmentation and basic analysis part, (2) model parameter estimation part and (3) model parameter tuning part (Fig. 1). Two cameras are used to capture images and graphical hand model describing recognition results and relating services are provide to user.

In the first part, hand region is segmented and the first analysis is conducted. Then, hand-model parameters are roughly estimated using database. Finally, fine modification of parameters is processed and response according to the recognition result.

This system recognizes 50 hand-signals. Those are selected from frequently used real driving hand-signals. They are composed with 32 hand-poses and 5 hand-trajectories.

The focal point of our discussion is hand-pose recognition and it is treated in the second part – model parameter estimation part. The discussion on hand region segmentation and trajectory recognition is beyond the scope of the present paper and dealt in [1].

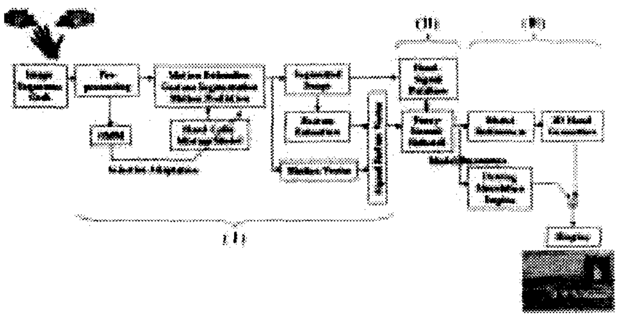

Fig.1 Overall flow of the target system

a. Hand region segmentation and first analysis
b. Model parameter estimation with database
c. Fine tuning and display the result

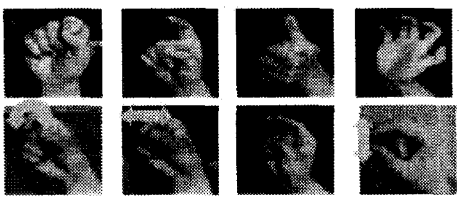

Fig.2 Hand-signal examples

The hand-pose recognition process relies on predefined database. Database is composed with representative hand poses and model parameters, that is, finger angles (Fig. 3).

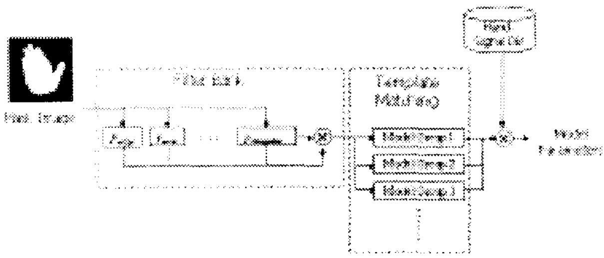

Fig.3 Rough model parameter estimation

Rough estimation is done by template matching between one segmented hand image from input and 32 representatives defined in the database.

Through the process, it showed good performance although there are fundamental problems caused by viewing direction and self-occlusion, not assuming highly restricted environment very small number of hand postures with limited viewing direction [4, 5, 6].

III. EXTENSION BY SUBGROUP-BASED CLASSIFICATION

The method showed in chapter II uses template matching for rough estimation. It has two problems as below.

First, template matching process is expensive in computation. Moreover each matching process requires multiple correlations to consider the rotation of hand. For one segmented hand region, template matching is repeated several times as many as the number of predefined hand pose.

The second problem is caused by using silhouette of hand for comparison. That means the result is weak in self-occlusion and the shape variation resulted by movement and rotation of hand in the image. If hand is viewed from different direction or self-occlusion is occurred, template matching result tends to be wrong.

In the case of second problem, it can be solved by adding extra features as a measure of similarity. But it needs to be done with attention to avoid computational complexity and extravagancy.

Here, we propose hand pose recognition method using decomposed approach with subgroup based scheme. First, input hand is decomposed into subgroups and then characteristic process is conducted depend on resulting subgroup (Fig. 4).

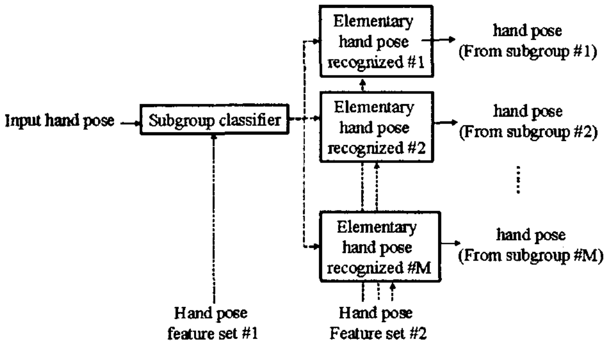

Fig.4 Recognition with decomposed approach (M: the number of hand pose subgroups)

This method not even reduces a lot of computation load but also provides the extendibility of hand poses defined in database. That is, instead of processing all the hand poses in the database, by classification into possible subgroup and calculates within the models in the subgroup, required time is decreased. Also the characterized process according to the classified subgroup is achieved. When new hand pose is added, the only thing to do is add the pos into similar subgroup and the whole correction is not required. More detailed process is described in the following chapter.

IV. EXPERMENTAL RESULTS

The total number of hand poses to recognize is 32. It is classified into 12 subgroups according to the similarity (Fig. 5). We used a simplified version similarity measure as (1) to reduce computational load. The normalized images are used. Normalization performed according to the primary axis and size of segmented hand region.

$$S = 1 + \frac{n(I_a \overline{\oplus} I_b) - n(I_a \oplus I_b)}{2n_T}, \quad i = 1, \cdots, 12 \quad (1)$$

I_a, I_b: Normalized hand images

$n(I)$: The Sum of pixels in image I

n_T: Total number of pixels in the whole image

$I_a \overline{\oplus} I_b$: Exclusive NOR between I_a and I_b

$I_a \oplus I_b$: Exclusive OR between I_a and I_b

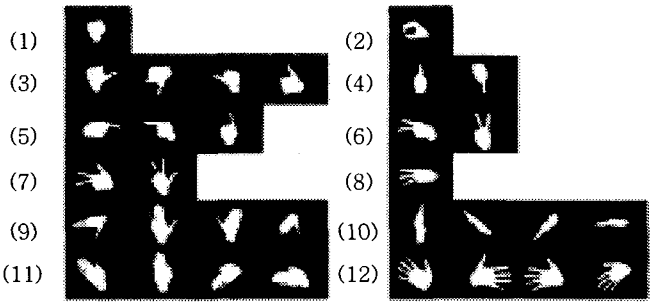

Fig. 5 Hand pose subgroups

The features used in the experiment are primary axis (θ), elongation (E) and onvexness (C) of hand region [7].

Above all, input are classified into proper subgroups and through individual process, recognition is performed. To define individual process related to each subgroup, we used C4.5 algorithm. Table 1 shows features used to classify each hand poses.

For the experiments, we used 6 datasets composed with 32 hand poses each extracted from 5 people. That is, for one hand pose, there are 30 data and therefore there are 960data in total. Among those data, we used 480 data for learning and 480 data for test. Table 2 shows experimental results.

Table 1. Features for individual recognition phase

Subgroup	Used features	The number of Elementary hand pose
(1)	E, C	1
(2)	E, C	1
(3)	θ, E	4
(4)	θ	2
(5)	θ, E	3
(6)	θ	2
(7)	θ	2
(8)	E, C	1
(9)	θ, E	4
(10)	θ	4
(11)	θ	4
(12)	θ, E	4

Table 2. Recognition rate

Hand pose	Recognition rate (%)	Hand pose	Recognition rate (%)	Hand pose	Recognition rate (%)
1	100	12	100	23	100
2	86.67	13	100	24	100
3	93.33	14	100	25	100
4	100	15	100	26	93.33
5	100	16	93.33	27	93.33
6	93.33	17	100	28	100
7	86.67	18	100	29	100
8	93.33	19	100	30	100
9	93.33	20	93.33	31	100
10	86.67	21	100	32	93.33
11	86.67	22	100	-	-
				합 계	96.46

V. CONCLUSION

In this paper, we used hierarchical approach to recognize hand pose. It showed remarkable structural simplicity, characterized process and extensibility with recognition rate of 96.46%.

Further works will be aim at addition of features and optimization of each recognition phase.

REFERENCES

[1] Hyoyoung Jang, Dae-Jin Kim, Jung-Bae Kim and Zeungnam Bien, "Study on Hand Pose Recognition Using Decomposed Approach with Subgroup-based scheme," Proc. of IEEK Summer Conference 2003, Vol.25, No. 1, pp. 215-218, 2002.

[2] G. L. Turbin, An introduction to matched filtering, IRE Trans. Inf. Theory, IT-6, No. 3, pp. 311-329, June 1960.

[3] W. K. Pratt, Digital Image Processing, 2nd ed., Wiley-Interscience, New York, 1991.

[4] R. Hamdan, F. Heitz, and L. Thoraval, Gesture localization and recognition using probabilistic visual learning, CVPR, vol. 2, pp. 98-103, 1999.

[5] M. Kohler, Special topics of gesture recognition applied in intelligent home environments, Proc. Of the Gesture Workshop, pp. 285-296, 1997.

[6] A. Utsumi and J. Ohya, Multiple-hand-gesture tracking using multiple cameras, CVPR, vol. 1, pp.473-478, 1999.

[7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.